FACULTY WORKING PAPER 92-0148

A Moving-Estimates Test for Parameter Stability and its Boundary-Crossing Probability

Chia-Shang J. Chu
Department of Economics
University of Southern California

Kurt Hornik
Institut für Statistik und Wahrscheinlichkeitstheorie
Technische Universität Wien

Chung-Ming Kuan
Department of Economics
University of Illinois at Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 92-0148

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

August 1992

A Moving-Estimates Test for Parameter Stability and its Boundary-Crossing Probability

Chia-Shang J. Chu

Kurt Hornik

Chung-Ming Kuan

A MOVING-ESTIMATES TEST FOR PARAMETER STABILITY AND ITS BOUNDARY-CROSSING PROBABILITY

Chia-Shang J. Chu
Department of Economics
University of Southern California

Kurt Hornik
Institut für Statistik und Wahrscheinlichkeitstheorie
Technische Universität Wien

Chung-Ming Kuan
Department of Economics
University of Illinois at Urbana-Champaign

This Draft: August 7, 1992

† We thank Roger Koenker for many useful discussions and comments.

Abstract

In this paper we propose a moving-estimates test for parameter stability and characterize its asymptotic null distribution analytically. The proposed test is based on fluctuation of moving estimates computed from a sequence of subsamples of the same size and is computationally simple. We obtain the limiting process of fluctuation of moving estimates using the functional central limit theorem and derive formulas representing the boundary-crossing probability of this limiting process, from which the asymptotic critical values of the proposed test are determined and tabulated. The proposed test is consistent for a general class of alternatives. Our simulation also shows that the proposed test has power superior to other competing tests when there are double structural changes.

JEL Classification Number: 211

Keywords: Moving Estimate, Fluctuation Test, Moving-Estimates Test, Wiener Process, Brownian Bridge, Structural Change, Boundary-Crossing Probability, Functional Central Limit Theorem.

1 Introduction

Testing for parameter constancy has recently been receiving much attention in the econometric literature. Contributions to this topic include the fluctuation (FL) test of Sen [15] and Ploberger, Krämer & Kontrus [14] and maximal Wald-type test of Hawkins [11], and maximal likelihood-ratio-type and Lagrange-Multiplier-type tests of Andrews [1]. In contrast with traditional testing procedures such as the Chow [5] test, a novel feature of the new testing results is that no prior knowledge of the location of change point is required. Under quite general conditions, these new test statistics have well defined asymptotic distributions in the space of continuous functions, and their critical values can be determined using the well-known boundary-crossing probability formulas associated with the Brownian bridge or tied-down Bessel process. These testing results have also been extended to models with trending regressors, e.g., Chu & White [6] and Hansen [10].

In this paper we propose a different test for parameter stability. Our test is similar to the FL test in spirit, in that the test statistic is determined by fluctuation of (a sequence of) parameter estimates. On the other hand, our test is implemented using *moving* estimates computed from a sequence of subsamples of the same size, whereas the FL test is based on *recursive* estimates calculated from a sequence of subsamples of increasing size. The proposed moving-estimates (ME) test is also different from the "homogeneity test" of Brown, Durbin & Evans [4] and the moving (rolling) t-tests of Banerjee, Lumsdaine & Stock [2]. Moving estimates have been used in econometrics mainly for tracking a time-varying system. Intuitively, moving estimates are more sensitive to parameter changes than recursive estimates, hence a smaller number of moving estimates suffices to detect parameter changes. The ME test is therefore computationally simple; in particular, the number of models which need to be estimated in the proposed test is about one half of that of the FL and maximal Wald-type tests. Our simulation also shows that the proposed test has power superior to other competing tests when there are certain double structural changes.

A further contributions of this paper is to characterize the asymptotic null distribution of the ME test statistics analytically. Under the null hypothesis that the true parameter is a constant, we find that fluctuations of moving estimates (in terms of their deviations from the full-sample average) converge weakly to the increments of a Brownian bridge. However, this limiting process is non-standard and, to the best of our knowledge, its boundary-crossing probabilities have not been investigated yet. We explicitly derive formulas to represent these probabilities, from which the asymptotic critical values of the ME test can be determined. Apart from its application to test of parameter constancy, the probability result we establish is *new* and interesting in its own right.

This paper is organized as follows. We introduce the ME test for the location model in section 2. The main asymptotic distribution result is derived in section 3. Extension to the multiple regression model is discussed in section 4. We report simulation results in section 5. Section 6 summarizes the paper. All proofs are deferred to the Appendix.

2 The Moving-Estimates Test

Consider the location model

$$y_i = \mu_i + \epsilon_i, \qquad i = 1, 2, \ldots, T.$$

We assume that $\{\epsilon_i\}$ is a sequence of random variables obeying a functional central limit theorem (FCLT), i.e., the normalized partial sum of ϵ_i converges weakly to the standard Wiener process:

$$\left(\frac{1}{\sigma\sqrt{T}} \sum_{i=1}^{[Tt]} \epsilon_i, 0 \leq t \leq 1 \right) \Rightarrow W \tag{1}$$

as $T \rightarrow \infty$, where

$$\sigma^2 = \lim_{T\rightarrow\infty} \frac{1}{T} \mathbb{E}\Big[\Big(\sum_{i=1}^{T} \epsilon_i\Big)^2\Big],$$

$[Tt]$ is the integer part of Tt, $\Rightarrow$ denotes weak convergence (of the associated probability measures), and W is the standard Wiener process. For more details about weak convergence and FCLT we refer to Billingsley [3]. We note that (1) holds under fairly general conditions and that (1) remains valid when σ is replaced by a consistent estimator $\hat{\sigma}$; see e.g., Wooldridge & White [16].

Under the null hypothesis that the parameter is a constant, $\mu_i = \mu_0$ for all i. In what follows we call a subsample a "window" (of the whole sample), and $\hat{\sigma}$ denotes an estimator of σ. We recall that the FL test computes the recursive estimates of μ_0 from a sequence of growing windows:

$$\hat{\mu}_k = \frac{1}{k} \sum_{i=1}^{k} y_i, \quad k = 1, \cdots, T, \tag{2}$$

and the test statistic is based on the differences between $\hat{\mu}_k$ and the full-sample average $\hat{\mu}_T$ under suitable normalization:

$$FL_T = \max_{1 \leq k \leq T} \frac{k}{\hat{\sigma}\sqrt{T}} |\hat{\mu}_k - \hat{\mu}_T|. \tag{3}$$

For given h $(0 \leq h \leq 1)$, the moving estimates of μ_0 are computed from windows, each with $[Th]$ observations, moving across the whole sample. That is,

$$\tilde{\mu}_{k,h} = \frac{1}{[Th]} \sum_{i=k+1}^{k+[Th]} y_i, \quad k = 0, \cdots, T - [Th]. \tag{4}$$

Note that moving estimates (4) are computed using only the most recent $[Th]$ observations in the window, and the information from the distant past (depending on the window size) is excluded gradually as windows move forward. On the other hand, recursive estimates (2) are obtained from growing windows so that no past information is discarded. Hence, moving estimates usually contain "purer" information and are more sensitive to parameter changes than recursive estimates.

Under the null hypothesis that the true parameter is a constant, moving estimates should not fluctuate too much and should be close to $\hat{\mu}_T$. The proposed ME test for parameter stability is therefore based on the differences between $\tilde{\mu}_{k,h}$ and $\hat{\mu}_T$. Typically, we are interested in the two-sided alternative: $\mu_i \neq \mu_0$ for some i. In some applications, it is possible that only an one-sided alternative (e.g., $\mu_i > \mu_0$ for some i) concerns us. This situation may arise when one has prior belief that the mean μ might have changed in only one direction or when the error of accepting the null hypothesis under $\mu_i < \mu_0$ is of no practical importance. For example, let y be the ratio of defective products of a production line. A quality control manager is only interested in whether there is a significant increase of this ratio; a decrease of the defective ratio is not practically relevant. The ME statistics for one- and two-sided alternatives are, respectively,

$$ME^{+}_{T,h} = \max_{k=0,\cdots,T-[Th]} \frac{[Th]}{\hat{\sigma}\sqrt{T}} (\tilde{\mu}_{k,h} - \hat{\mu}_T), \tag{5}$$

$$ME_{T,h} = \max_{k=0,\cdots,T-[Th]} \frac{[Th]}{\hat{\sigma}\sqrt{T}} |\tilde{\mu}_{k,h} - \hat{\mu}_T|. \tag{6}$$

Let S_T denote the piecewise constant interpolation of

$$S_T\left(\frac{k}{T}\right) = \frac{1}{\hat{\sigma}\sqrt{T}} \sum_{i=1}^{k} \epsilon_i,$$

so that $S_T(t) = (\hat{\sigma}\sqrt{T})^{-1} \sum_{i=1}^{[Tt]} \epsilon_i$. S_T is a process in $D([0,1])$, the space of functions that are right continuous with left-hand limits on $[0,1]$, and $S_T \Rightarrow W$ by the FCLT (1). Also let S_T^0 denote the tied-down process given by

$$S_T^0(t) = S_T(t) - \frac{[Tt]}{T} S_T(1)$$

which is also in $D([0,1])$ with jumps at k/T, $1 \leq k \leq T$. Note that $S_T^0(0) = S_T^0(1) = 0$ and $S_T^0 \Rightarrow W^0$, where W^0 is a Brownian bridge. Also note that S_T^0 attains its extrema at one of the jump points k/T. Under the null hypothesis,

$$\begin{aligned}
&\frac{[Th]}{\hat{\sigma}\sqrt{T}} (\tilde{\mu}_{k,h} - \hat{\mu}_T) \\
&= \frac{1}{\hat{\sigma}\sqrt{T}} \left(\sum_{i=1}^{k+[Th]} \epsilon_i - \sum_{i=1}^{k} \epsilon_i \right) - \frac{[Th]}{T} \frac{1}{\hat{\sigma}\sqrt{T}} \sum_{i=1}^{T} \epsilon_i \\
&= S_T\left(\frac{k+[Th]}{T}\right) - S_T\left(\frac{k}{T}\right) - \frac{[Th]}{T} S_T(1) \\
&= \left[S_T\left(\frac{k+[Th]}{T}\right) - \frac{k+[Th]}{T} S_T(1)\right] - \left[S_T\left(\frac{k}{T}\right) - \frac{k}{T} S_T(1)\right] \\
&= S_T^0\left(\frac{k}{T} + h_T\right) - S_T^0\left(\frac{k}{T}\right).
\end{aligned} \tag{7}$$

where $h_T = [Th]/T$. The associated empirical ME process

$$\begin{aligned}
M_{T,h}(t) &= \frac{[Th]}{\hat{\sigma}\sqrt{T}} (\tilde{\mu}_{[Nt],h} - \hat{\mu}_T) \\
&= S_T^0\left(\frac{[Nt]}{T} + h_T\right) - S_T^0\left(\frac{[Nt]}{T}\right), \qquad 0 \leq t \leq 1-h,
\end{aligned} \tag{8}$$

is the piecewise constant interpolation on $[0, 1-h]$ of (7) with interpolation nodes k/N, $0 \leq k \leq T - [Th]$, where $N = (T - [Th])/(1-h)$. Observe that $[Nt]/T \to t$ and $h_T \to h$ as $T \to \infty$. With a little extra work, the following result follows from the continuous mapping theorem.

Theorem 2.1 *Assume that the FCLT (1) holds. Then under the null hypothesis, if $\hat{\sigma}$ is consistent for σ, we have*

$$M_{T,h} \Rightarrow M_h,$$

where $M_h(t) = W^0(t+h) - W^0(t)$ for $0 \leq t \leq 1-h$. In particular,

$$\begin{aligned} ME^{+}_{T,h} &\Rightarrow \max_{0 \leq t \leq 1-h} M_h(t), \\ ME_{T,h} &\Rightarrow \max_{0 \leq t \leq 1-h} |M_h(t)|. \end{aligned}$$

This result says that the empirical ME process converges in distribution to the increments of a Brownian bridge. The probability that the limiting process M_h crosses the boundaries α and β ($\alpha < \beta$) at least once on $[0, 1-h]$ is

$$\mathbb{P}\{M_h(t) > \beta \text{ or } M_h(t) < \alpha \text{ for some } 0 \leq t \leq 1-h\}.$$

Clearly, for $\alpha = -\infty$ and $\alpha = -\beta$, the boundary-crossing probabilities determine the asymptotic critical values of the ME test statistics for one- and two-sided alternatives, respectively. These probabilities are evaluated in the next section.

There is an obvious dilemma in choosing the window size h. When h is large, moving estimates are relatively more precise under the null hypothesis. If the true parameter follows two regimes, a large h implies that a smaller number of estimates can be used to detect parameter instability, and many moving estimates will incorporate data from two regimes. If h is small, on the other hand, we have more (but relatively imprecise) estimates to detect parameter changes, but the detected structural instability might just be a consequence of sample variation. In view of this trade-off, a natural choice is $h = 1/2$. Thus, each moving estimate is calculated using $[T/2]$ observations. We also note that the proposed ME test computes at most $T - [T/2] + 1$ moving estimates, in contrast with the FL test which may have to compute T recursive estimates and other competing tests such as the maximal Wald test which must estimate two models for each hypothetical change point. Hence, the proposed ME test is computationally simpler. Given current computing capability, computational simplicity is by no means our major concern. What is remarkable is that the proposed test, which uses much less estimates, can outperform other competing tests when there are multiple structural changes; see Section 5 for simulation results.

3 Asymptotic Null Distributions

A major challenge of this research is that, to the best of our knowledge, there is no known boundary-crossing probability result for the limiting M_h process. This is different from the situation encountered in [1] and [14], where the boundary-crossing probability results of the tied-down Bessel process and Brownian bridge are well known in literature. Nonetheless, we show that the asymptotic distribution of the ME statistic can be characterized analytically as well.

The process M_h is clearly Gaussian with mean zero and continuous paths. For $0 \leq s \leq t \leq 1-h$, we have

$$\begin{aligned}
&\text{cov}(M_h(s), M_h(t)) \\
&= \text{cov}(W^0(s+h), W^0(t+h)) - \text{cov}(W^0(s), W^0(t+h)) \\
&\qquad - \text{cov}(W^0(s+h), W^0(t)) + \text{cov}(W^0(s), W^0(t)) \\
&= [(s+h) - (s+h)(t+h)] - [s - s(t+h)] \\
&\qquad - [\min(s+h, t) - (s+h)t] + [s - st] \\
&= h(1-h) - \min(s+h, t) + s \\
&= h(1-h) - \min(h, t-s).
\end{aligned}$$

The covariance function of M_h is thus

$$\text{cov}(M_h(s), M_h(t)) = \sigma_h^2 - \min(h, |t-s|), \tag{9}$$

where $\sigma_h^2 = h(1-h)$ is the variance of the M_h process. In particular, $M_h(0)$ is normal with mean zero and variance σ_h^2.

Conditional on $M_h(0)$, we can represent the M_h process in terms of a Brownian bridge by rescaling the time parameter, as shown in the following lemma.

Lemma 3.1 *Let $h \geq 1/2$. Given $M_h(0) = m$,*

$$M_h =^d \left((1 - t/\sigma_h^2)m + 2\sigma_h W^0(t/2\sigma_h^2),\ 0 \leq t \leq 1-h\right)$$

where $=^d$ denotes equality in distribution of two random processes.

By Lemma 3.1, the conditional probability that $M_h(t)$ remains within the constant boundaries α and β on $[0, 1-h]$ can be written as the probability that a Brownian bridge stays within two linear boundaries. That is,

$$\begin{aligned}
&p_h(\alpha, \beta; m)\\
&\quad := \mathbb{P}\{\alpha \le M_h(t) \le \beta \text{ for all } 0 \le t \le 1-h | M_h(0) = m\}\\
&\quad = \mathbb{P}\{\alpha \le (1 - t/\sigma_h^2)m + 2\sigma_h W^0(t/2\sigma_h^2) \le \beta \text{ for all } 0 \le t \le 1-h\}\\
&\quad = \mathbb{P}\{\alpha \le (1-2u)m + 2\sigma_h W^0(u) \le \beta \text{ for all } 0 \le u \le 1/2h\},
\end{aligned}$$

where the last equality follows by letting $u = t/2\sigma_h^2$. We now confine ourselves to the case that $h = 1/2$. Then, $\sigma_h^2 = 1/4$ and

$$\begin{aligned}
&p_{1/2}(\alpha, \beta; m)\\
&\quad = \mathbb{P}\{(\alpha - m) + 2um \le W^0(u) \le (\beta - m) + 2um \text{ for all } 0 \le u \le 1\}.
\end{aligned} \tag{10}$$

It is easy to show that $(1+t)W^0(t/(1+t))$ is in fact a Wiener process. Now, if $a, b, c, d > 0$, by letting $u = t/(1+t)$ we can write

$$\begin{aligned}
&\mathbb{P}\{-a(1-u) - bu \le W^0(u) \le c(1-u) + du \text{ for all } 0 \le u \le 1\}\\
&\quad = \mathbb{P}\left\{-a - bt \le (1+t)W^0(t/(1+t)) \le c + dt \text{ for all } t \ge 0\right\}\\
&\quad = \mathbb{P}\{-a - bt \le W(t) \le c + dt \text{ for all } t \ge 0\}\\
&\quad = 1 - \sum_{k=1}^{\infty} (e^{-2A_k} + e^{-2B_k} - e^{-2C_k} - e^{-2D_k}),
\end{aligned} \tag{11}$$

where

$$\begin{aligned}
A_k &= k^2 cd + (k-1)^2 ab + k(k-1)(ad + bc),\\
B_k &= (k-1)^2 cd + k^2 ab + k(k-1)(ad + bc),\\
C_k &= k^2(ab + cd) + k(k-1)ad + k(k+1)bc,\\
D_k &= k^2(ab + cd) + k(k+1)ad + k(k-1)bc;
\end{aligned}$$

the last equality of (11) is a well-known formula due to Doob [7, p. 398]. For

$$a = -(\alpha - m), \quad b = -(\alpha + m), \quad c = \beta - m, \quad d = \beta + m,$$

the left-hand side of (11) is just the right-hand side of (10). We thus have:

Lemma 3.2 *Conditional on* $M_{1/2}(0) = m$,

$$p_{1/2}(\alpha,\beta;m) = 1 - \sum_{k=1}^{\infty}(e^{-2A_k} + e^{-2B_k} - e^{-2C_k} - e^{-2D_k}),$$

where

$$\begin{aligned} A_k &= (k\beta - (k-1)\alpha)^2 - m^2, \\ B_k &= ((k-1)\beta - k\alpha)^2 - m^2, \\ C_k &= (k(\beta - \alpha) - m)^2 - m^2, \\ D_k &= (k(\beta - \alpha) + m)^2 - m^2, \end{aligned}$$

for $\alpha \leq 0$, $\beta \geq 0$, *and* $|m| \leq \min(|\alpha|, \beta)$; *otherwise*, $p_{1/2}(\alpha,\beta;m) = 0$.

For $\alpha \leq 0 \leq \beta$, the unconditional probability that $M_{1/2}(t)$ stays within the boundaries α and β is given by

$$\begin{aligned} &\mathbb{P}\{\alpha \leq M_{1/2}(t) \leq \beta \text{ for all } 0 \leq t \leq 1/2\} \\ &= \sqrt{2/\pi}\int_{|m|\leq\min(|\alpha|,\beta)} p_{1/2}(\alpha,\beta;m)\, e^{-2m^2}\, dm, \end{aligned} \tag{12}$$

where $\sqrt{2/\pi}e^{-2m^2}$ is the density of $M_{1/2}(0)$, which is normally distributed with mean zero and variance 1/4. Of particular interest to us is of course the case that $\alpha = -\infty$ and $\alpha = -\beta$. By evaluating the integral in (12), we finally obtain:

Theorem 3.3 *For all* $\beta > 0$,

$$\mathbb{P}\{M_{1/2}(t) \leq \beta \textit{ for all } 0 \leq t \leq 1/2\} = 2\Phi(2\beta) - 1 - 4\beta\phi(2\beta) \tag{13}$$

and

$$\begin{aligned} &\mathbb{P}\{|M_{1/2}(t)| \leq \beta \textit{ for all } 0 \leq t \leq 1/2\} \\ &= 1 - 8\beta\sum_{k=1}^{\infty}\phi(2(2k-1)\beta) &(14) \\ &= 2\sum_{k=1}^{\infty}(-1)^{k+1}e^{-k^2\pi^2/8\beta^2}, &(15) \end{aligned}$$

where Φ *and* ϕ *are the distribution and density functions of the standard normal random variable.*

In light of Theorem 2.1 and Theorem 3.3, the asymptotic null distributions of the ME tests follow immediately.

Corollary 3.4 *Assume that the FCLT* (1) *holds. Then under the null hypothesis, if $\hat{\sigma}$ is consistent for σ, we have*

$$\lim_{T\to\infty} \mathbb{P}\{ME^{+}_{T,1/2} \leq \beta\} = 2\Phi(2\beta) - 1 - 4\beta\phi(2\beta),$$

$$\lim_{T\to\infty} \mathbb{P}\{ME_{T,1/2} \leq \beta\} = 2\sum_{k=1}^{\infty}(-1)^{k+1}e^{-k^2\pi^2/8\beta^2}.$$

Some asymptotic critical values β are summarized in Table 1 below. We stress that these values are *not* results of model-specific simulation.

Table 1: The Critical Values of the ME Test for the Location Model.

ME	Probabilities			
Tests	0.90	0.95	0.975	0.99
One-Sided	1.25014	1.39774	1.52876	1.68411
Two-Sided	1.37506	1.51151	1.63408	1.78082

Note: The critical values are solved numerically using *Mathematica*. For two-sided tests, β's are solved from (15) with 10 terms in the summation, even though we notice that the terms with $k \geq 5$ are virtually negligible.

Remarks:

(1) It is straightforward to verify that for $\beta > 0$,

$$\mathbb{P}\{M_{1/2}(t) \geq -\beta \text{ for all } 0 \leq t \leq 1/2\} = 2\Phi(2\beta) - 1 - 4\beta\phi(2\beta)$$

which is the same as (13).

(2) If β goes to 0, it is clear that the probability in (15) converges to 0, as it ought to be. However, it may seem strange that as β goes to 0, (13) also tends to zero. Observe that $M_{1/2}(0) = -M_{1/2}(1/2)$. Hence, $\max_{0\leq t\leq 1/2} M_{1/2}(t)$ is always non-negative and $\min_{0\leq t\leq 1/2} M_{1/2}(t)$ is always non-positive.

(3) The probability of $M_{1/2}(t)$ crossing both boundaries on [0, 1/2] is not negligible. For example, the one-sided boundary 1.39774 at probability 0.95 cannot be used as the two-sided boundary at probability 0.90.

(4) Following Durbin [8], we can approximate the first passage density q of $M_{1/2}(t)$ to a constant boundary β as $q(t) \simeq 4\beta\sqrt{2/\pi}e^{-2\beta^2}$. The probability that $M_{1/2}(t)$ crosses β (from below) at least once on [0, 1/2] is thus

$$\int_0^{1/2} q(t)\,dt \simeq 2\beta\sqrt{2/\pi}e^{-2\beta^2}.$$

It follows that the probability that $M_{1/2}(t)$ never crosses β on [0, 1/2] is

$$\mathbb{P}\{M_{1/2}(t) \leq \beta \text{ for all } 0 \leq t \leq 1/2\} \simeq 1 - 2\beta\sqrt{2/\pi}e^{-2\beta^2} = 1 - 4\beta\phi(2\beta), \tag{16}$$

which is greater than (13) unless $\beta = \infty$. Hence, for given probability, Durbin's approximation (16) must underestimate the correct one-sided boundary. For example, for probabilities equal to 0.95 and 0.975, (16) yields one-sided boundaries β as 1.37506 and 1.51151, respectively, which are smaller than those given in Table 1. Coincidentally, these values turn out to be the same (up to five decimal places) as the correct two-sided boundaries in Table 1 for probabilities equal to 0.90 and 0.95, respectively. To see this, note that the first term ($k = 1$) in the summation of (14) is just 2 times the tail probability $4\beta\phi(2\beta)$ in Durbin's approximation and that all other terms in (14) are virtually negligible.

4 Extension to Multiple Regression

The previous results can be extended easily to the multiple linear regression model. We now consider the model

$$y_i = x_i'\theta_i + \epsilon_i, \qquad i = 1, 2, \cdots, T,$$

where x_i is a $n \times 1$ vector. The null hypothesis is that $\theta_i = \theta_0$ for all i.

We assume that the double array $\{x_i\epsilon_i/\sqrt{T}\}$ satisfies the conditions of Corollary 4.2 of Wooldridge & White [16] so that a multivariate FCLT holds:

$$\left(\frac{1}{\sqrt{T}}\Sigma^{-1/2}\sum_{i=1}^{[Tt]} x_i\epsilon_i, 0 \leq t \leq 1\right) \Rightarrow \boldsymbol{W}, \tag{17}$$

where

$$\Sigma := \lim_{T\to\infty} \frac{1}{T}\,\mathbb{E}\Big[\Big(\sum_{i=1}^{T} x_i\epsilon_i\Big)\Big(\sum_{i=1}^{T} x_i\epsilon_i\Big)'\Big],$$

and $\boldsymbol{W}$ stands for the n-dimensional, standard Wiener process, cf. (1). To reduce technicality, we do not state the regularity conditions explicitly, but we note that Corollary 4.2 of [16] allows x_i and ϵ_i to be weakly dependent, heterogeneous random variables but *not* integrated of positive order; see also [13]. The limiting result in (17) again holds when Σ is replaced by a consistent estimator $\hat{\Sigma}_T$; for example, $\hat{\Sigma}_T$ may be a heteroskedasticity and autocorrelation consistent estimator, e.g., Newey & West [12]. In addition to the FCLT, we also assume that the following weak law of large numbers (WLLN)

$$\frac{1}{[Th]} \sum_{i=[Nt]+1}^{[Nt]+[Th]} x_i x_i' \longrightarrow^{p} Q, \tag{18}$$

holds uniformly in $0 \le t \le 1-h$ for given h, where as usual $N = (T-[Th])/(1-h)$, $\longrightarrow^{p}$ stands for convergence in probability and Q is a non-singular, non-stochastic $n \times n$ matrix. Note that if $t = 0$ and $h = 1$, (18) gives the standard WLLN. It is not too hard to show that if

$$\frac{1}{T}\sum_{i=1}^{T} x_i x_i' \longrightarrow Q$$

almost surely, then (18) holds uniformly in $0 \le t \le 1-h$. We omit the details.

For given h, the moving OLS estimates are

$$\tilde{\theta}_{k,h} = \left(\sum_{i=k+1}^{k+[Th]} x_i x_i'\right)^{-1} \left(\sum_{i=k+1}^{k+[Th]} x_i y_i\right), \qquad k = 0,\cdots,T-[Th]. \tag{19}$$

Let $\hat{\theta}_T$ be the standard (full-sample) OLS estimator, $Q_T = T^{-1}\sum_{i=1}^{T} x_i x_i'$, and $\hat{D}_T = Q_T^{-1}\hat{\Sigma}_T Q_T^{-1}$ such that $\hat{D}_T^{-1/2} = \hat{\Sigma}_T^{-1/2} Q_T$. Here, $\hat{\Sigma}_T$ is again an estimator of Σ. For an n-dimensional vector V, let $\|V\| = \max_{j=1,\cdots,n}|V_j|$ be the maximum norm of V, where V_j is the j-th element of V. The ME test statistic for the multiple regression model is defined as

$$ME_{T,h} = \max_{k=0,\cdots,T-[Th]} \frac{[Th]}{\sqrt{T}} \|\hat{D}_T^{-1/2}(\tilde{\theta}_{k,h} - \hat{\theta}_T)\|. \tag{20}$$

Analogous to the notations used in Section 2 let

$$S_T(t) = \frac{1}{\sqrt{T}} \hat{\Sigma}_T^{-1/2} \sum_{i=1}^{[Tt]} x_i \epsilon_i$$

and let S_T^0 denote the tied-down process given by

$$S_T^0(t) = S_T(t) - \frac{[Th]}{T} S_T(1).$$

Under the null hypothesis, we have

$$\begin{aligned}
&\frac{[Th]}{\sqrt{T}} \hat{D}_T^{-1/2} (\tilde{\theta}_{k,h} - \hat{\theta}_T) \\
&= \frac{1}{\sqrt{T}} \hat{D}_T^{-1/2} \left(\frac{1}{[Th]} \sum_{i=k+1}^{k+[Th]} x_i x_i' \right)^{-1} \left(\sum_{i=1}^{k+[Th]} x_i \epsilon_i - \sum_{i=1}^{k} x_i \epsilon_i \right) \\
&\qquad - \frac{[Th]}{T} \hat{D}_T^{-1/2} \left(\frac{1}{T} \sum_{i=1}^{T} x_i x_i' \right)^{-1} \left(\frac{1}{\sqrt{T}} \sum_{i=1}^{T} x_i \epsilon_i \right) \\
&= \frac{1}{\sqrt{T}} \hat{\Sigma}_T^{-1/2} \left[Q_T \left(\frac{1}{[Th]} \sum_{i=k+1}^{k+[Th]} x_i x_i' \right)^{-1} - I \right] \left(\sum_{i=1}^{k+[Th]} x_i \epsilon_i - \sum_{i=1}^{k} x_i \epsilon_i \right) \\
&\qquad + S_T\left(\frac{k+[Th]}{T} \right) - S_T\left(\frac{k}{T} \right) - \frac{[Th]}{T} S_T(1). \qquad (21)
\end{aligned}$$

It is readily seen that the first term on the right-hand side of (21) is vanishingly small under WLLN (18) and the last three terms are $S_T^0(k/T + h_T) - S_T^0(k/T)$. Let the multivariate empirical ME process $M_{T,h}$ be the piecewise constant interpolation of (21), cf. (8), and let W^0 be an n-dimensional Brownian bridge. The following result is a multivariate extension of Theorem 2.1.

Theorem 4.1 *Assume that the FCLT* (17) *and WLLN* (18) *hold. Then under the null hypothesis, if $\hat{\Sigma}$ is consistent for Σ, we have*

$$M_{T,h} \Rightarrow M_h,$$

where $M_h(t) = W^0(t+h) - W^0(t)$ for $0 \le t \le 1-h$. In particular,

$$ME_{T,h} \Rightarrow \max_{0 \le t \le 1-h} \|M_h(t)\|.$$

For $h = 1/2$, we obtain from Theorem 3.3 that:

Theorem 4.2 *Assume that the FCLT* (17) *and WLLN* (18) *hold. Then under the null hypothesis, if* $\hat{\Sigma}$ *is consistent for* Σ, *we have*

$$\lim_{T\to\infty} \mathbb{P}\{ME_{T,1/2} \leq \beta\} = \left(2\sum_{k=1}^{\infty}(-1)^{k+1}e^{-k^2\pi^2/8\beta^2}\right)^n,$$

where n *is the number of parameters in the model.*

In Table 2 we summarize the asymptotic critical values for the regression model with 1 to 10 variables at various probabilities. Clearly, the critical values for $n = 1$ are just those given in the last row of Table 1.

[Table 2 About Here]

We now consider a general class of alternatives:

$$\theta_i = \theta_0 + T^{-\delta} g(i/T), \tag{22}$$

where $g\colon [0,1] \to \mathbb{R}^n$ is a (non-constant) vector-valued function of bounded variation on $[0,1]$. If $\delta = 0$, (22) is a global alternative; if $\delta = 1/2$, (22) characterizes a sequence of local alternatives. g may e.g. be a step function to represent multiple structural changes or a continuous function to represent smooth or periodic parameter changes.

Theorem 4.3 *Assume that the FCLT* (17) *and WLLN* (18) *hold. Then under the alternative* (22), *if* $\hat{\Sigma}_T^{-1}$ *is* $O_p(1)$, *we have*

$$T^{\delta-1/2}ME_{T,h} = \max_{0\leq t\leq 1-h}\left\|\hat{\Sigma}_T^{-1/2}(QL_h g(t) + T^{\delta-1/2}\Sigma^{1/2}M_{T,h}(t))\right\| + o_p(1), \tag{23}$$

where

$$L_h g(t) = \int_t^{t+h} g(u)\,du - h\int_0^1 g(u)\,du.$$

Notice that under the alternative, $\hat{\Sigma}_T$ is not necessarily consistent for Σ; in particular, under global alternatives ($\delta = 0$), it is not consistent in standard cases (see further below).

As the derivative of the function $L_h g$ at t is $g(t+h) - g(t)$, $L_h g$ is nonzero provided that g is not periodic with period h. Conversely, if g has period h, and $1/h$ is an integer, then

$$L_h g(t) = \int_t^{t+h} g(u)\,du - h\int_0^1 g(u)\,du = \int_0^h g(u)\,du - h\frac{1}{h}\int_0^h g(u)\,du = 0.$$

Hence, if in addition to the conditions of theorem 4.3, $\hat{\Sigma}_T$ is also $O_p(1)$, $0 \le \delta < 1/2$ and g is not periodic with period h, the right-hand side of (23) is bounded away from zero (in probability), and the ME test statistic grows at rate $T^{1/2-\delta}$; therefore, the ME test is consistent against such sequences of alternatives. On the other hand, the test is not consistent if g has period h and $1/h$ is an integer.

More definite results on the asymptotics of the ME test under alternatives (22) can be given under suitable assumptions on the structure of Σ. Suppose that $\{x_t\}$ is a sequence of suitably mixing random variables and that $\{\epsilon_t\}$ is a sequence of i.i.d. random variables with mean zero and variance σ^2 which is independent of $\{x_t\}$, as assumed in Ploberger, Krämer, & Kontrus [14, p. 308]. Given these conditions, $\Sigma = \sigma^2 Q$, and a natural estimate is

$$\hat{\Sigma}_T = \hat{\sigma}^2 Q_T, \qquad \hat{\sigma}^2 = \frac{1}{T}\sum_{i=1}^{T}(y_i - x_i'\hat{\theta}_T)^2. \tag{24}$$

It can easily be shown that

$$\hat{\sigma}^2 \to^p \sigma_\delta^2 = \begin{cases} \sigma^2, & 0 < \delta \le \frac{1}{2}, \\ \sigma^2 + \int_0^1 \left(g(u) - \int_0^1 g(v)\,dv\right) Q \left(g(u) - \int_0^1 g(v)\,dv\right)' du, & \delta = 0; \end{cases}$$

notice that $\sigma_0^2 > \sigma^2$ unless g in constant. Hence, we have the following.

Corollary 4.4 *Assume that the conditions in* [14, p. 308] *hold and that* $\hat{\Sigma}$ *is given by* (24). *Then under the alternatives* (22) *with* $\delta = 1/2$,

$$ME_{T,h} \Rightarrow \max_{0\le t\le 1-h} \left\| M_h(t) + \sigma^{-1}Q^{1/2}L_h g(t)\right\|;$$

if $0 \le \delta < 1/2$,

$$T^{\delta-1/2} ME_{T,h} \to^p \sigma_\delta^{-1} \max_{0\le t\le 1-h} \|Q^{1/2}L_h g(t)\|.$$

Remark: In view of the remarks after Theorem 4.3, if $\delta = 1/2$ and g is periodic with period h and $1/h$ is an integer,

$$ME_{T,h} \Rightarrow \max_{0 \le t \le 1-h} ||M_h(t)||,$$

which is identical to the limit under the null hypothesis. Thus, the ME test has only trivial local power against such alternatives.

5 Test Performance and Simulation

In this section we report simulation results and analyze the test performance under single structural change. The competing tests we consider are the FL and maximal Lagrange Multiplier (LM) tests. We do not consider the maximal likelihood ratio and Wald tests because they are based on the same ingredients as the FL test.

We first simulate empirical test sizes. The data y_i are generated from i.i.d. $N(2, 1)$, and the samples are $T = 100$, 200, 300 and 500. Note that in the ME test the window sizes are $h = 1/2$. In the LM test, it is required that for the hypothetical change points $[Tc]$, c must take values in some subset of $[0, 1]$ whose closure lies in $(0, 1)$, Andrews [1]. In our simulation, we take $c \in [0.1, 0.9]$. The variance estimate we used in these tests is the standard one:

$$\hat{\sigma}^2 = \sum_{i=1}^{T} (y_i - \hat{\mu}_T)^2 / T,$$

where $\hat{\mu}_T$ is the full-sample average. These results are summarized in Table 3. It can be seen that the ME test has quite reasonable empirical sizes; in particular, it has the least finite-sample size distortion when $T = 500$.

[Table 3 About Here]

We also simulate the power performance based on a single structural change and double structural changes. The data generating process (DGP) for a single structural change is

$$y_i = \begin{cases} \mu_0 + \epsilon_i, & i = 1, \cdots, [T\lambda], \\ \mu_0 + \Delta + \epsilon_i, & i = [T\lambda] + 1, \cdots, T, \end{cases} \tag{25}$$

where $\mu_0 = 2$ and ϵ_i are i.i.d. $N(0,1)$. We consider $\Delta = 0.4$ & 0.2 and $\lambda = 0.1, 0.2, \cdots, 0.9$ for samples $T = 150$ and 300. Because the power performance is symmetric in λ, we only report the results for $\Delta = 0.4$ and λ from 0.1 to 0.5 in Table 3. A complete result is available from the authors upon request. It can be seen that the ME test is less favorable than the other two tests. Nevertheless, the ME test performs quite well when sample is 300 and change points are close to the center of the sample.

[Table 4 About Here]

The power deficiency of the ME test in the case of a single structural change is in fact to be expected. In the Appendix we show that, under the alternative (25),

$$\frac{1}{\sqrt{T}} ME_{T,1/2} \to^p \frac{1}{2} \min(\lambda, 1-\lambda) \frac{|\Delta|}{\sigma_0}, \tag{26}$$

and

$$\frac{1}{\sqrt{T}} FL_T \to^p \lambda(1-\lambda) \frac{|\Delta|}{\sigma_0}, \tag{27}$$

where σ_0 is defined after (24). Without loss of generality, set $|\Delta| = 1$ and $\sigma_0 = 1$. Thus, (27) is a parabola (opening downward) with vertex (1/2,1/4) and zero points (0,0) and (1,0), and (26) is the linear interpolation of (0,0), (1/2,1/4) and (1,0) which is inside the parabola (27). Note that the asymptotic critical values of the ME test are greater than those of the FL test, hence the FL test rejects whenever the ME test rejects. It follows from (26) and (27) that the power of the FL test dominates that of the ME test for all possible change points λ, at least asymptotically, and both power functions attain the maximum at $\lambda = 1/2$. This is indeed compatible with our simulation results.

However, the ME test has power advantage when there are double structural changes. For double structural changes, the DGP is

$$y_i = \begin{cases} \mu_0 + \epsilon_i, & i = 1, \cdots, [T\lambda_1], \\ \mu_0 + \Delta_1 + \epsilon_i, & i = [T\lambda_1] + 1, \cdots, [T\lambda_2], \\ \mu_0 + \Delta_2 + \epsilon_i, & i = [T\lambda_2] + 1, \cdots, T. \end{cases} \tag{28}$$

We consider two cases: $\Delta_1 = 0.4$ & $\Delta_2 = 0.1$ and $\Delta_1 = 0.4$ & $\Delta_2 = 0$, both with $\mu_0 = 2$ and the break points $\lambda_1 = 0.3$ and $\lambda_2 = 0.4, 0.5, \cdots, 0.9$. The samples are $T = 150$ and 300. These results are collected in Table 5 and 6. It can be seen that the ME test dominates the competing tests in *all* cases considered, and this dominance becomes very significant when the second break point occurs at and after 0.6. Further, the ME test has the highest power when $\lambda_1 = 0.3$ and $\lambda_2 = 0.8$. In this case, the difference of two break points is equal to the size of windows in the test. Also note that the FL and LM tests perform quite similarly. It is not surprising to see that if two break points are very close, e.g., $\lambda_1 = 0.3$ and $\lambda_2 = 0.4$, all the tests have relatively low power. The DGPs we simulate here are not unreasonable in reality. When there is a shock in economy, economic variables may behave differently and return to previous level after certain period of time.

[Tables 5 and 6 About Here]

6 Summary

In this paper a ME test for parameter stability is proposed and its asymptotic distribution is characterized analytically. We derive formulas representing the boundary-crossing probabilities of moving estimates and solve for the correct critical values. These critical values are tabulated for future reference. When there are double structural changes, our simulation results show that the proposed test compares favorably with the FL and LM tests. Thus, the ME test can complement other tests for parameter stability. There are some limitations of our results, however. First, our results do not apply to models with integrated and trending regressors. Second, the boundary-crossing probability formula is derived only for $h = 1/2$. Extension to a general h appears to be very challenging. The result of Durbin [8] may provide an approximate solution.

Appendix

Lemma A *Let X_T be a sequence of random processes in $D([0,1])^k$ converging in distribution (with respect to the Skorohod topology) to a random process X in $C([0,1])^k$ (i.e., the limiting process has continuous paths). Further, let $0 < h_T < 1$ be a sequence converging to $0 < h < 1$, and let $\kappa_T : [0, 1-h] \to [0, 1-h_T]$ be a sequence of maps such that $\sup_{0\leq t\leq 1-h} |\kappa_T(t) - t|$ tends to zero. Then, if Z_T is the random process on $D([0,1-h])^k$ given by*

$$Z_T(t) = X_T(\kappa_T(t) + h_T) - X_T(\kappa_T(t)),$$

we have $Z_T \Rightarrow Z$, where for $0 \leq t \leq 1-h$, $Z(t) = X(t+h) - X(t)$.

Proof: For a function f in $D([0,1])^k$, put

$$\omega_f(\delta) = \max_{0\leq s,t\leq 1; |s-t|\leq\delta} |f(s) - f(t)|.$$

Also, let $\delta_T = \max\left(|h_T - h|, \sup_{0\leq t\leq 1-h} |\kappa_T(t) - t|\right)$ and let $Z_T^*(t) = X_T(t+h) - X_T(t)$. Then $\delta_T \to 0$ and the continuous mapping theorem yields that $Z_T^* \Rightarrow Z$. As

$$\begin{aligned} &|Z_T(t) - Z_T^*(t)| \\ \leq\ & |X_T(\kappa_T(t) + h_T) - X_T(t+h)| + |X_T(\kappa_T(t)) - X_T(t)| \\ \leq\ & 2\omega_{X_T}(2\delta_T), \end{aligned}$$

we have $\sup_{0\leq t\leq 1-h} |Z_T(t) - Z_T^*(t)| \leq 2\omega_{X_T}(2\delta_T) = o_p(1)$ because $X_T \Rightarrow X$ and X has continuous sample paths. It follows from Theorem 4.1 of Billingsley (1968) that Z_T has the same weak limit as Z_T^*, whence $Z_T \Rightarrow Z$. □

Proof of Theorem 2.1: We apply Lemma A. Setting $X_T = S_T^0$, $X = W^0$, and $\kappa_T(t) = [Nt]/T$. Note that for $0 \leq t \leq 1-h$, $0 \leq \kappa_T(t) \leq 1 - h_T$, where $h_T = [Th]/T$. Clearly, $\sup_{0\leq t\leq 1-h} |\kappa_T(t) - t| \to 0$. It follows from (8) and Lemma A that $M_{T,h} \Rightarrow M_h$. Since $M_{T,h}$ is piecewise constant, it reaches its extrema at one of the jumps. Thus, from (6) we have

$$\begin{aligned} ME_{T,h}^+ &= \max_{0\leq t\leq 1-h} M_{T,h}(t), \\ ME_{T,h} &= \max_{0\leq t\leq 1-h} |M_{T,h}(t)|. \end{aligned}$$

The limits of the ME statistics now follow from the first assertion and the continuous mapping theorem. □

Proof of Lemma 3.1: If $h \geq 1/2$, the covariance function of $M_h(t)$ becomes

$$\mathrm{cov}(M_h(s), M_h(t)) = \sigma_h^2 - |t - s|$$

by (9). Thus, for $0 \leq s \leq t \leq 1 - h$, $[M_h(0), M_h(s), M_h(t)]'$ has a normal distribution with mean zero and covariance matrix

$$\begin{bmatrix} \sigma_h^2 & \sigma_h^2 - s & \sigma_h^2 - t \\ \sigma_h^2 - s & \sigma_h^2 & \sigma_h^2 - t + s \\ \sigma_h^2 - t & \sigma_h^2 - t + s & \sigma_h^2 \end{bmatrix}.$$

Using the standard regression formulas for the conditional normal distribution, we find that given $M_h(0) = m$, M_h is a Gaussian process with mean function

$$\begin{aligned} &\mathbb{E}(M_h(t)|M_h(0) = m) \\ &= \mathbb{E}M_h(t) + \frac{\mathrm{cov}(M_h(t), M_h(0))}{\mathrm{var}(M_h(0))}(m - \mathbb{E}M_h(0)) \\ &= (1 - t/\sigma_h^2)m \end{aligned} \qquad (29)$$

and covariance function

$$\begin{aligned} &\mathrm{cov}(M_h(s), M_h(t)|M_h(0) = m) \\ &= \mathrm{cov}(M_h(s), M_h(t)) - \frac{\mathrm{cov}(M_h(s), M_h(0))\,\mathrm{cov}(M_h(0), M_h(t))}{\mathrm{var}(M_h(0))} \\ &= \sigma_h^2 - t + s - \frac{(\sigma_h^2 - s)(\sigma_h^2 - t)}{\sigma_h^2} \\ &= 2s - st/\sigma_h^2. \end{aligned} \qquad (30)$$

Thus, given $M_h(0) = m$,

$$\left(\frac{1}{2\sigma_h}(M_h(2\sigma_h^2 u) - (1 - 2u)m),\ 0 \leq u \leq 1/2h\right),$$

is Gaussian with continuous sample paths, mean zero, and covariance function

$$\frac{1}{4\sigma_h^2}\left[2 \cdot 2\sigma_h^2 u - \frac{2\sigma_h^2 u \cdot 2\sigma_h^2 v}{\sigma_h^2}\right] = u - uv, \qquad 0 \leq u \leq v \leq 1/2h,$$

by (29) and (30), i.e., a Brownian bridge on $[0, 1/2h]$. Hence, given $M_h(0) = m$, we have

$$\left(M_h(t) - (1 - t/\sigma_h^2)m,\ 0 \leq t \leq 1 - h\right) =^d \left(2\sigma_h W^0(t/2\sigma_h^2),\ 0 \leq t \leq 1 - h\right). \quad \square$$

Proof of Lemma 3.2: The conclusion is obvious from the text. We only have to verify the ranges of the parameters stated in the Lemma. In view of (11), a and c must be non-negative because the standard Wiener process starts at 0. This implies that $\alpha \leq m \leq \beta$. Also, a Wiener process changes its sign infinitely often with probability one, hence b and d are necessarily non-negative. It follows that $-(\alpha - m) - (\alpha + m) \geq 0$, i.e., $\alpha \leq 0$. Similarly, $\beta \geq 0$. Thus, $p_{1/2}(\alpha, \beta; m) = 0$ unless $\alpha \leq 0$, $\beta \geq 0$, and $|m| \leq \min(|\alpha|, \beta)$. □

Proof of Theorem 3.3: For the one-sided case, $\alpha = -\infty$, and all A_k, B_k, C_k, D_k approach infinity except $A_1 = \beta^2 - m^2$. Hence,

$$\begin{aligned}
&\mathbb{P}\{M_{1/2}(t) \leq \beta \text{ for all } 0 \leq t \leq 1/2\} \\
&= \sqrt{2/\pi} \int_{|m| \leq \beta} (1 - e^{-2(\beta^2 - m^2)})\, e^{-2m^2}\, dm \\
&= \sqrt{2/\pi} \int_{|m| \leq \beta} (e^{-2m^2} - e^{-2\beta^2})\, dm \\
&= (2\pi)^{-1/2} \left(\int_{|u| \leq 2\beta} e^{-u^2/2}\, du - 4\beta e^{-(2\beta)^2/2} \right) \\
&= 2\Phi(2\beta) - 1 - 4\beta\phi(2\beta).
\end{aligned}$$

For $\alpha = -\beta$, we have

$$\begin{aligned}
A_k = B_k &= (2k-1)^2\beta^2 - m^2, \\
C_k &= (m - 2k\beta)^2 - m^2, \\
D_k &= (m + 2k\beta)^2 - m^2.
\end{aligned}$$

It is readily verified that

$$\begin{aligned}
\sqrt{2/\pi} \int_{-\beta}^{\beta} e^{-2(m-y)^2}\, dm &= (2\pi)^{-1/2} \int_{-2(\beta+y)}^{2(\beta-y)} e^{-u^2/2}\, du \\
&= \Phi(2(\beta - y)) - \Phi(-2(\beta + y))
\end{aligned}$$

which coincides for $y = \pm 2k\beta$ and gives $\Phi(2(2k+1)\beta) - \Phi(2(2k-1)\beta)$. From (12),

$$\begin{aligned}
&\mathbb{P}\{|M_{1/2}(t)| \leq \beta \text{ for all } 0 \leq t \leq 1/2\} \\
&= \sqrt{2/\pi} \int_{-\beta}^{\beta} p_{1/2}(-\beta, \beta; m)\, e^{-2m^2}\, dm \\
&= 2\Phi(2\beta) - 1 - 2\sqrt{2/\pi}\, 2\beta \sum_{k=1}^{\infty} e^{-2(2k-1)^2\beta^2}
\end{aligned}$$

$$+2\sum_{k=1}^{\infty}\Phi(2(2k+1)\beta)-\Phi(2(2k-1)\beta)$$

$$= \quad 2\Phi(2\beta)-1-4\sqrt{2/\pi}\,\beta\sum_{k=1}^{\infty}e^{-2(2k-1)^2\beta^2}+2(1-\Phi(2\beta))$$

$$= \quad 1-8\beta\sum_{k=1}^{\infty}\phi(2(2k-1)\beta).$$

This establishes (14). To prove that (14) and (15) are equal, first observe that

$$8\beta\sum_{k=1}^{\infty}\phi(2(2k-1)\beta)=4\beta\sum_{k=-\infty}^{\infty}\phi(2(2k-1)\beta)=\beta\sqrt{\frac{8}{\pi}}\sum_{k=-\infty}^{\infty}e^{-2(2k-1)^2\beta^2}.$$

Then using the alternative form of the Poisson summation formula in Feller [9, p.632] with the density of the standard normal random variable $f(x)=(2\pi)^{-1/2}e^{-x^2/2}$ (such that the characteristic function $\varphi(a)=e^{-a^2/2}$), $\zeta=\lambda$ and $s=0$, we have

$$\sum_{k=-\infty}^{\infty}e^{-(2k-1)^2\lambda^2/2}=\sqrt{\frac{\pi}{2\lambda^2}}\sum_{k=-\infty}^{\infty}(-1)^k e^{-k^2\pi^2/2\lambda^2}.$$

For $\lambda=2\beta$ we obtain

$$\beta\sqrt{\frac{8}{\pi}}\sum_{k=-\infty}^{\infty}e^{-2(2k-1)^2\beta^2} = \beta\sqrt{\frac{8}{\pi}}\sqrt{\frac{\pi}{8\beta^2}}\sum_{k=-\infty}^{\infty}(-1)^k e^{-k^2\pi^2/8\beta^2}$$

$$= \sum_{k=-\infty}^{\infty}(-1)^k e^{-k^2\pi^2/8\beta^2}$$

and finally

$$1-8\beta\sum_{k=1}^{\infty}\phi(2(2k-1)\beta) = 1-\sum_{k=-\infty}^{\infty}(-1)^k e^{-k^2\pi^2/8\beta^2}$$

$$= 2\sum_{k=1}^{\infty}(-1)^{k+1}e^{-k^2\pi^2/8\beta^2}. \quad \square$$

Proof of Corollary 3.4: ¿From Theorems 2.1 and 3.3 we have

$$\lim_{T\to\infty}\mathbb{P}\{ME^{+}_{T,1/2}\le\beta\}$$

$$= \mathbb{P}\{M_{1/2}(t)\le\beta \text{ for all } 0\le t\le 1/2\}$$

$$= 2\Phi(2\beta)-1-4\beta\varphi(2\beta).$$

The second assertion follows similarly. □

Proof of Theorem 4.1: As the square bracket term in (21):

$$Q_T \left(\frac{1}{[Th]} \sum_{i=[Nt]+1}^{[Nt]+[Th]} x_i x_i' \right)^{-1} - I$$

converges to zero in probability uniformly in t by the WLLN (18),

$$\boldsymbol{M}_{T,h}(t) - (\boldsymbol{S}_T^0([Nt]/T + h_T) - \boldsymbol{S}_T^0([Nt]/T)) \longrightarrow^p 0$$

uniformly in t. Hence, it suffices to consider the interpolation of $\boldsymbol{S}_T^0([Nt]/T + h_T) - \boldsymbol{S}_T^0([Nt]/T)$ for $0 \leq t \leq 1 - h$ as in (8). The rest of the proof is therefore the same as the proof of Theorem 2.1. □

Proof of Theorem 4.2: Note that the elements W_j^0, $j = 1, \ldots, n$ of $\boldsymbol{W}^0$ are mutually independent, univariate Brownian bridges. By Theorem 4.1, we obtain

$$\begin{aligned}
&\lim_{T\to\infty} \mathbb{P}\{ME_{T,1/2} \leq \beta\} \\
&= \mathbb{P}\{\|\boldsymbol{W}^0(t + 1/2) - \boldsymbol{W}^0(t)\| \leq \beta \text{ for all } 0 \leq t \leq 1/2\} \\
&= \mathbb{P}\{|W_j^0(t + 1/2) - W_j^0(t)| \leq \beta \text{ for all } j = 1, \cdots, n,\ 0 \leq t \leq 1/2\} \\
&= \left(\mathbb{P}\{|W_j^0(t + 1/2) - W_j^0(t)| \leq \beta \text{ for all } 0 \leq t \leq 1/2\}\right)^n \\
&= \left(\mathbb{P}\{|M_{1/2}(t)| \leq \beta \text{ for all } 0 \leq t \leq 1/2\}\right)^n.
\end{aligned}$$

The assertion now follows from Theorem 3.3. □

Proof of Theorem 4.3: Under the alternative (22),

$$\begin{aligned}
\hat{\theta}_T &= \theta_0 + \left(\frac{1}{T}\sum_{i=1}^T x_i x_i'\right)^{-1} \left(\frac{1}{T^{1+\delta}} \sum_{i=1}^T x_i x_i' g(i/T) + \frac{1}{T}\sum_{i=1}^T x_i \epsilon_i\right) \\
&=: \theta_0 + Q_T^{-1} \left(\frac{1}{T^{1+\delta}} R_T + \frac{1}{T}\sum_{i=1}^T x_i \epsilon_i\right),
\end{aligned}$$

and the moving OLS estimates are

$$\begin{aligned}
\tilde{\theta}_{k,h} &= \theta_0 + \left(\frac{1}{[Th]} \sum_{i=k+1}^{k+[Th]} x_i x_i'\right)^{-1} \\
&\quad \left(\frac{1}{T^{1+\delta} h_T} \sum_{i=k+1}^{k+[Th]} x_i x_i' g(i/T) + \frac{1}{[Th]} \sum_{i=k+1}^{k+[Th]} x_i \epsilon_i\right)
\end{aligned}$$

$$=: \quad \theta_0 + Q_{k,T}^{-1}\left(\frac{1}{T^{1+\delta}h_T}R_{k,T} + \frac{1}{[Th]}\sum_{i=k+1}^{k+[Th]} x_i\epsilon_i\right).$$

Hence,

$$\begin{aligned}
&T^{\delta-1/2}ME_{T,h}\\
&= \max_{0\le t\le 1-h} T^{\delta}h_T\|\hat{D}_T^{-1/2}(\tilde{\theta}_{[Nt],h} - \hat{\theta}_T)\|\\
&= \max_{0\le t\le 1-h}\Bigg\|\hat{\Sigma}_T^{-1/2}Q_T\Bigg[Q_{[Nt],T}^{-1}\left(\frac{1}{T}R_{[Nt],T} + T^{\delta-1/2}\frac{1}{\sqrt{T}}\sum_{i=[Nt]+1}^{[Nt]+[Th]} x_i\epsilon_i\right)\\
&\qquad - Q_T^{-1}\left(h_T\frac{1}{T}R_T + T^{\delta-1/2}h_T\frac{1}{\sqrt{T}}\sum_{i=1}^{T} x_i\epsilon_i\right)\Bigg]\Bigg\|.
\end{aligned}$$

Provided the WLLN (18) holds and that g is of bounded variation on $[0,1]$, it can be shown that

$$\frac{1}{T}R_{[Nt],T} = \frac{1}{T}\sum_{i=[Nt]+1}^{[Nt]+[Th]} x_ix_i'g(i/T) \to^p Q\int_t^{t+h} g(u)\,du$$

uniformly on $[0,1-h]$ and that

$$\frac{1}{T}R_T = \frac{1}{T}\sum_{i=1}^{T} x_ix_i'g(i/T) \to^p Q\int_0^1 g(u)\,du.$$

Hence,

$$\begin{aligned}
&\hat{\Sigma}_T^{-1/2}Q_T\left[Q_{[Nt],T}^{-1}\left(\frac{1}{T}R_{[Nt],T}\right) - Q_T^{-1}h_T\left(\frac{1}{T}R_T\right)\right]\\
&= \hat{\Sigma}_T^{-1/2}Q\left[Q^{-1}Q\int_t^{t+h} g(u)\,du - hQ^{-1}Q\int_0^1 g(u)\,du\right] + o_p(1)\\
&= \hat{\Sigma}_T^{-1/2}QL_hg(t) + o_p(1)
\end{aligned}$$

uniformly on $[0,1-h]$. As clearly

$$Q_TQ_{[Nt],T}^{-1}\frac{1}{\sqrt{T}}\sum_{i=[Nt]+1}^{[Nt]+[Th]} x_i\epsilon_i - h_T\frac{1}{\sqrt{T}}\sum_{i=1}^{T} x_i\epsilon_i = \Sigma^{1/2}M_{T,h}(t) + o_p(1)$$

uniformly on $[0,1-h]$ and $\hat{\Sigma}_T^{-1}$ is $O_p(1)$ by assumption, we obtain

$$\begin{aligned}
&T^{\delta-1/2}ME_{T,h}\\
&= \max_{0\le t\le 1-h}\left\|\hat{\Sigma}_T^{-1/2}\left(QL_hg(t) + T^{\delta-1/2}(\Sigma^{1/2}M_{T,h}(t) + o_p(1))\right)\right\|\\
&= \max_{0\le t\le 1-h}\left\|\hat{\Sigma}_T^{-1/2}(QL_hg(t) + T^{\delta-1/2}\Sigma^{1/2}M_{T,h}(t))\right\| + o_p(1). \quad \square
\end{aligned}$$

Proof of Corollary 4.4: The conclusions follow easily from Theorem 4.3 and (24). □

Proof of Equations (26) and (27): In the location model with a single structural break, we have from the second assertion of Corollary 4.4 that

$$\frac{1}{\sqrt{T}} ME_{T,1/2} \to^{p} \max_{0 \le t \le 1/2} \frac{1}{\sigma_0} \left| \int_t^{t+1/2} g(u)\, du - \frac{1}{2} \int_0^1 g(u)\, du \right|,$$

where $g(u) = \Delta$ for $u > \lambda$ and $g(u) = 0$ otherwise. It is easy to see that

$$\left| \int_t^{t+1/2} g(u)\, du - \frac{1}{2} \int_0^1 g(u)\, du \right| = \begin{cases} (1-\lambda)|\Delta|/2, & t + 1/2 \le \lambda, \\ |(t - \lambda/2)\Delta|, & t \le \lambda < t + 1/2, \\ \lambda|\Delta|/2, & t > \lambda. \end{cases}$$

Hence,

$$\max_{0 \le t \le 1/2} \frac{1}{\sigma_0} \left| \int_t^{t+1/2} g(u)\, du - \frac{1}{2} \int_0^1 g(u)\, du \right| = \frac{1}{2} \min(\lambda, 1 - \lambda) \frac{|\Delta|}{\sigma_0}.$$

This proves (26). Let $(c)^+$ denote the positive part of c. Under the alternative (25), the recursive estimates are

$$\hat{\mu}_j = \mu_0 + \left(\frac{j - [T\lambda]}{j} \right)^+ \Delta + \frac{1}{j} \sum_{i=1}^{j} \epsilon_i,$$

Again, it suffices to consider the behavior of the deterministic component of these estimates. We then have

$$\begin{aligned} \max_j \frac{1}{\sqrt{T}} \frac{j}{\hat{\sigma}\sqrt{T}} |\hat{\mu}_j - \hat{\mu}_T| &= \frac{[T\lambda]}{\hat{\sigma} T} \left(\frac{T - [T\lambda]}{T} \right) |\Delta| + o_p(1) \\ &\to^{p} \lambda(1 - \lambda) \frac{|\Delta|}{\sigma_0}, \end{aligned}$$

establishing (27). □

References

1. Andrews, D. W. K. Tests for parameter instability and structural change with unknown change points. Cowles Foundation Discussion Paper 943, Yale University, 1990.

2. Banerjee, A., R. L. Lumsdaine, & J. H. Stock. Recursive and sequential tests of the unit root and trend break hypotheses: Theory and international evidence. Working Paper, Kennedy School of Government, Harvard University, 1990.

3. Billingsley, P. *Convergence of Probability Measures.* New York: Wiley, 1968.

4. Brown, R. L., J. Durbin, & J. M. Evans. Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society*, Series B, **37** (1975): 149–163.

5. Chow, G. C. Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, **28** (1960): 591–605.

6. Chu, C.-S. J., & H. White. Testing for structural change in some simple time-series models. *Journal of Business and Economic Statistics*, to appear.

7. Doob, J. L. Heuristic approach to the Kolmogorov-Smirnov theorems. *Annals of Mathematical Statistics*, **20** (1949): 393–403.

8. Durbin, J. The first-passage density of a continuous Gaussian process to a general boundary. *Journal of Applied Probability*, **22** (1985): 99–122.

9. Feller, W. *An Introduction to Probability Theory and Its Applications*, Volume II, Second edition. New York: Wiley, 1971.

10. Hansen, B. E. Testing for structural change of unknown form in models with non-stationary regressors. Working Paper, Department of Economics, University of Rochester, 1991.

11. Hawkins, D. L. A test for a change point in a parametric model based on a maximal Wald-type statistic. *Sankhyā: Indian Journal of Statistics*, **49** (1987): 368–376.

12. Newey, W. K., & K. D. West. A simple, positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, **55** (1987): 703–708.

13. Phillips, P. C. B., & S. N. Durlauf. Multiple time series regression with integrated processes. *Review of Economic Studies*, **53** (1986): 473–495.

14. Ploberger, W., W. Krämer, & K. Kontrus. A new test for structural stability in the linear regression model. *Journal of Econometrics*, **40** (1989): 307–318.

15. Sen, P. K. Asymptotic theory of some tests for a possible change in the regression slope occurring at an unknown time point. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **52** (1980): 203–218.

16. Wooldridge, J. M., & H. White. Some invariance principles and central limit theorems for dependent heterogeneous processes. *Econometric Theory*, **4** (1988): 210–230.

Table 2: The Critical Values of the ME Test for the Multiple Regression Model.

Number of Parameters	Probabilities			
	0.90	0.95	0.975	0.99
n=1	1.37506	1.51151	1.63408	1.78082
n=2	1.50667	1.63193	1.74546	1.88269
n=3	1.57852	1.69814	1.80711	1.93951
n=4	1.62747	1.74345	1.84947	1.97871
n=5	1.66437	1.77772	1.88160	2.00854
n=6	1.69387	1.80519	1.90740	2.03255
n=7	1.71838	1.82805	1.92891	2.05261
n=8	1.73931	1.84760	1.94734	2.06980
n=9	1.75753	1.86465	1.96342	2.08483
n=10	1.77366	1.87976	1.97769	2.09819

Note: The critical values are solved numerically from the formula of Theorem 4.2 with 10 terms in the summation.

Table 3: Size Simulation of the ME, FL, and LM Tests.

Tests	Sample Sizes			
	T=100	T=200	T=300	T=500
ME	7.0%	7.5%	8.2%	9.2%
FL	7.7%	8.1%	8.4%	8.5%
LM	7.6%	7.9%	8.2%	8.3%

Note: Observations are generated from i.i.d. $N(2,1)$. The critical values of the ME, FL, and LM tests at 10% significance level are 1.375, 1.224, and 7.73, respectively. The number of replications is 2,500.

Table 4: Power Simulation under a Single Structural Change: $\Delta = 0.4$.

λ	$T = 150$			$T = 300$		
	ME	FL	LM	ME	FL	LM
0.1	11.6%	14.0%	20.7%	17.3%	23.6%	39.0%
0.2	21.1%	33.6%	38.6%	39.4%	62.0%	68.3%
0.3	35.1%	52.9%	49.1%	63.4%	83.1%	80.8%
0.4	47.0%	63.0%	55.3%	80.1%	91.1%	86.4%
0.5	54.9%	66.5%	57.4%	86.9%	92.4%	87.7%

Note: The DGP is based on (25) in the text with $\mu_0 = 2$, $\Delta = 0.4$, and ϵ_i i.i.d. $N(0,1)$. The number of replications is 2,500.

Table 5: Power Simulation under Double Structural Changes: $\Delta_1 = 0.4$ & $\Delta_2 = 0.1$.

λ_2	$T = 150$			$T = 300$		
	ME	FL	LM	ME	FL	LM
0.4	11.5%	11.7%	11.2%	14.5%	14.1%	15.6%
	14.1%	13.5%	13.2%	17.6%	15.8%	13.1%
0.5	17.7%	15.2%	14.4%	30.0%	24.1%	24.6%
	21.8%	17.6%	17.9%	36.3%	30.5%	28.4%
0.6	30.0%	21.0%	20.0%	51.0%	37.0%	35.0%
	35.3%	23.1%	23.6%	55.0%	41.0%	38.0%
0.7	39.6%	26.0%	26.5%	68.7%	48.9%	48.5%
	45.3%	28.6%	28.8%	71.3%	53.1%	51.5%
0.8	49.1%	31.7%	33.9%	80.0%	61.2%	61.4%
	54.8%	36.6%	39.1%	90.0%	62.7%	62.8%
0.9	42.9%	40.4%	39.6%	74.4%	69.4%	68.9%
	50.2%	45.8%	45.4%	78.9%	74.6%	72.7%

Note: In each cell, the first and second numbers are the empirical power based on asymptotic and empirical critical values, respectively. The DGP is (28) in the text with $\mu_0 = 2$, $\Delta_1 = 0.4$ & $\Delta_2 = 0.1$, $\lambda_1 = 0.3$, and ϵ_i i.i.d. $N(0,1)$. The number of replications is 2,500.

Table 6: Power Simulation under Double Structural Changes: $\Delta_1 = 0.4$ & $\Delta_2 = 0$.

λ_2	$T = 150$			$T = 300$		
	ME	FL	LM	ME	FL	LM
0.4	11.0%	10.9%	09.4%	17.6%	16.9%	13.9%
	16.0%	13.9%	12.5%	21.4%	20.0%	14.2%
0.5	21.3%	17.0%	14.8%	42.1%	30.7%	25.7%
	27.2%	22.0%	20.0%	43.8%	34.5%	29.0%
0.6	36.0%	24.0%	21.0%	65.0%	45.0%	40.0%
	42.8%	27.8%	26.1%	68.0%	50.0%	44.0%
0.7	46.3%	28.4%	28.3%	79.6%	51.7%	51.6%
	54.3%	30.0%	31.9%	82.4%	57.6%	56.2%
0.8	53.2%	29.7%	32.0%	87.2%	58.8%	63.2%
	60.3%	32.5%	37.6%	93.3%	60.8%	64.8%
0.9	47.9%	36.8%	38.2%	79.8%	67.9%	69.8%
	55.9%	41.5%	43.9%	80.7%	68.7%	71.4%

Note: This simulation uses the same DGP as that of Table 5 except that $\Delta_2 = 0$.